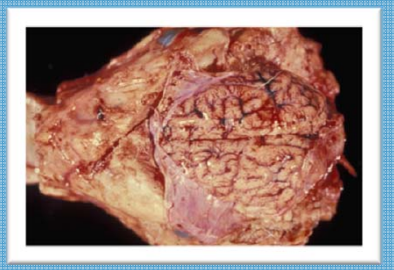
Using genomic sequences to develop new diagnostic methods for Glässer's disease

Kate Howell

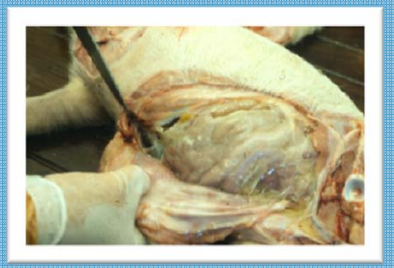
PhD Student
Department of Veterinary Medicine
Supervisors: Dr Tucker & Prof. Maskell

Introduction

- Glässer's disease is a systemic disease of pigs caused by *Haemophilus parasuis.*



- Disease caused by *H. parasuis* has been on the increase, with larger outbreaks seen since the adoption of higher health herds and multi-site practices.



- The only tests currently available in the UK are able to detect presence/absence of the bacterium.
- Serotyping is available in Europe but is expensive and can only type 80% of isolates.

- Vaccines are available but only effective against 2/15 serotypes.

Objectives

Main Objective:

To develop new molecular diagnostics against *H. parasuis* that would be able to predict the serotype of a strain and predict the likelihood of a strain causing disease.

-> Investigation of genetic determinants of serotype

-> Identification of putative virulence factors comparing whole genome sequence data of disease-associated and non- associated strains.

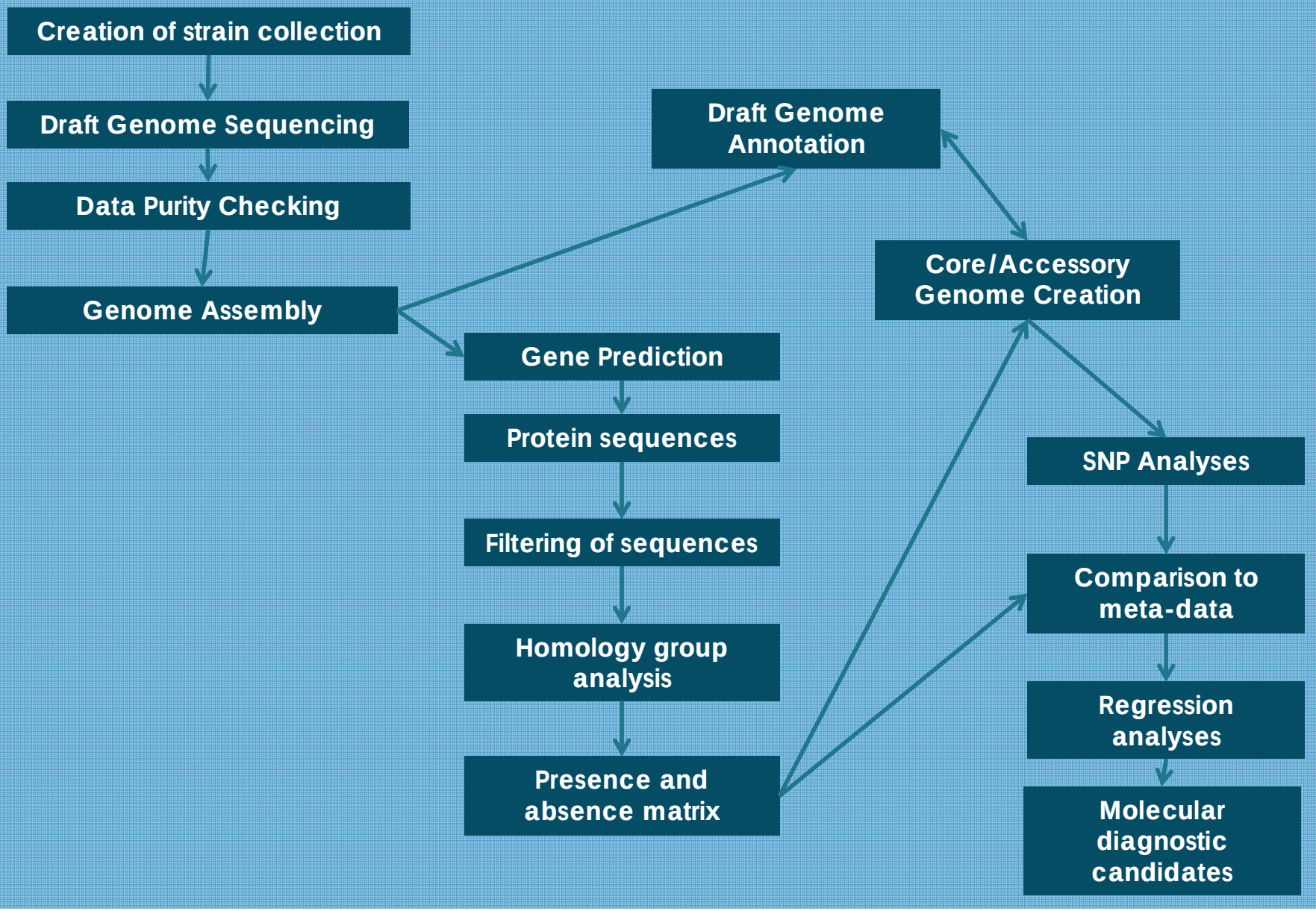
-> Development of molecular tests, e.g. PCR, that rapidly detect *H. parasuis* in field samples and indicate the likely importance of the isolate for the disease

-> Validation of the newly developed PCR tests using field samples collected from commercial herds.

Industry Focus

- HPS has a significant impact of disease on mortality, productivity and welfare of pigs globally.
- The proposed diagnostic would fill a hole in the market by offering molecular serotyping in the UK and information on the chance of a strain causing disease.
- It would be a cheap and fast diagnostic to allow fast turn around of results.

Materials and Methods



Homology Groups

A single homology group represents a "unique protein" found amongst all of the proteins predicted from our genome sequences.

Example 1:

One protein that is only found in 5/200 strains should be represented by its own homology group.

Example 2:

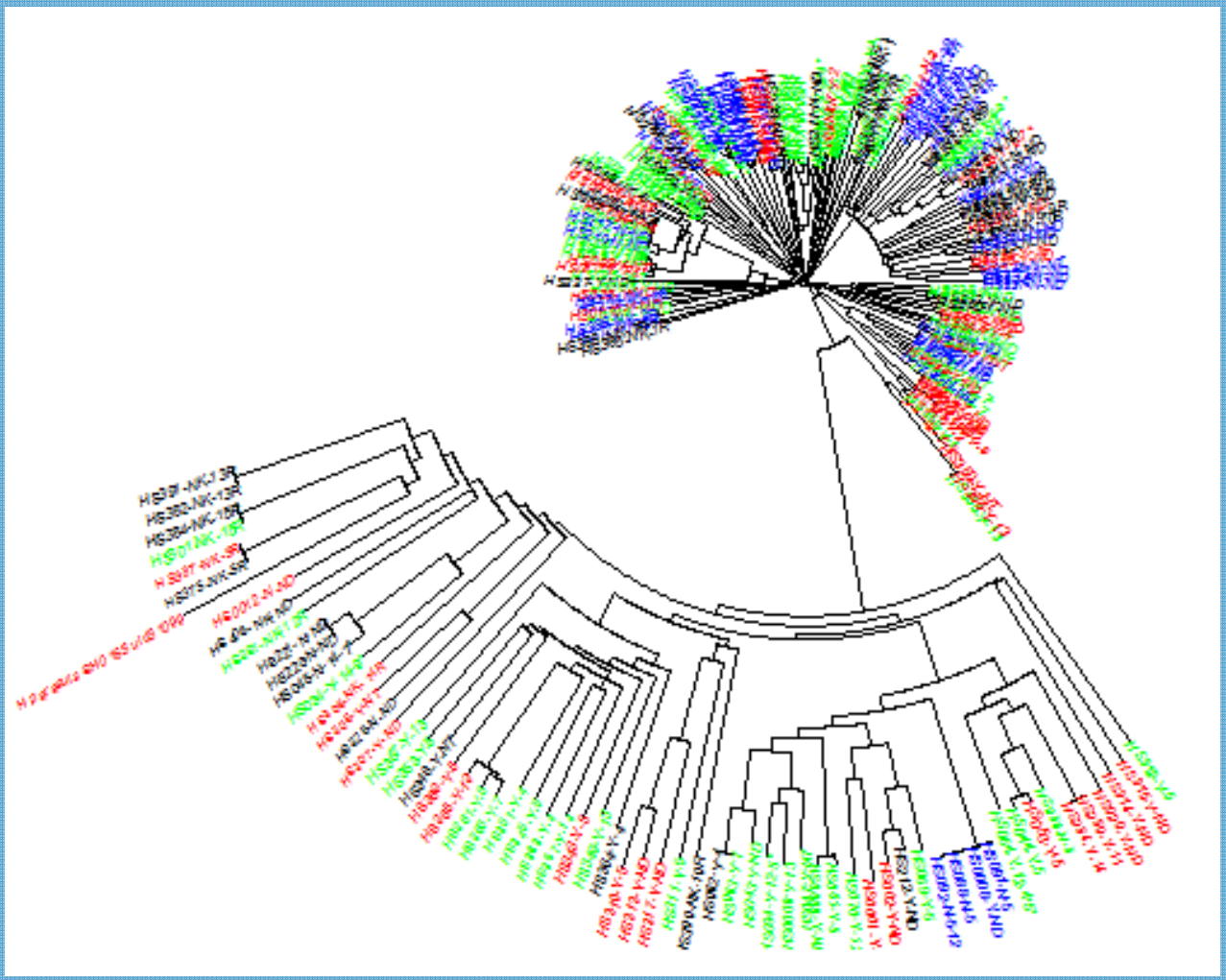
An essential protein found in every strain, would be another homology group.

Results – Core Genome

Non-clinical

Systemic

Respiratory

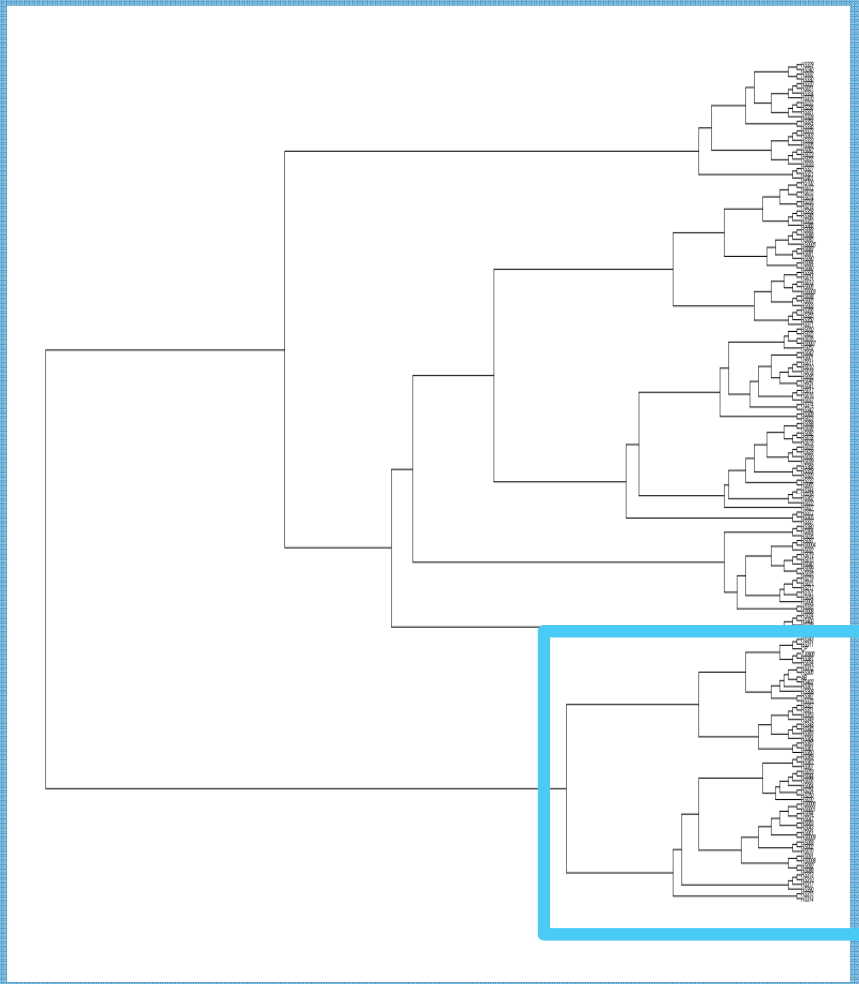


Results – Core vs. Accessory

Comparison of the Core and Accessory Trees shows similar shapes.

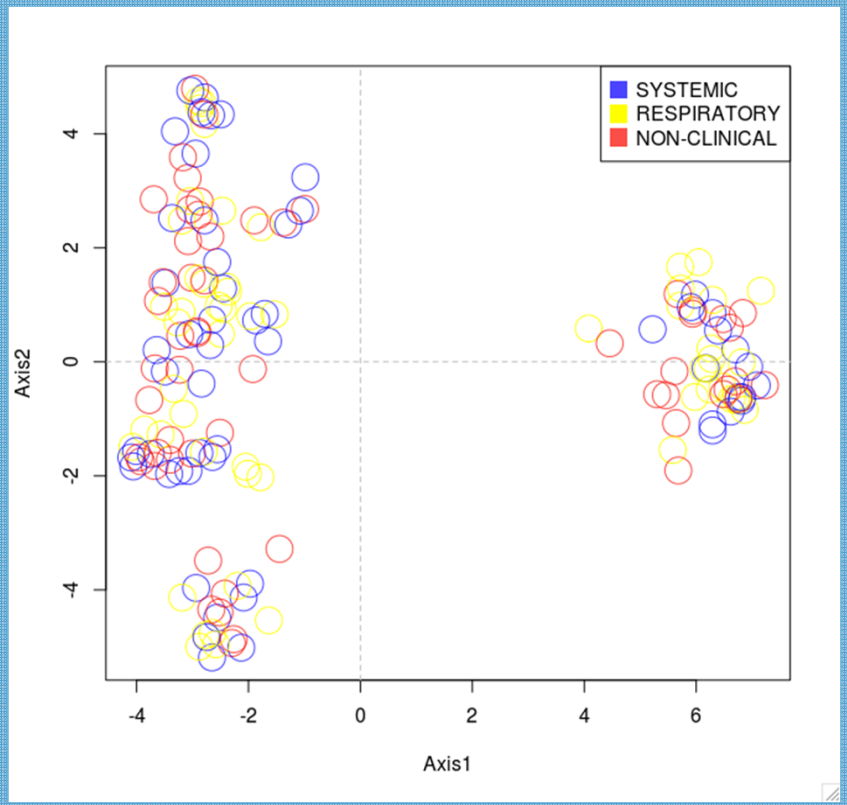
CORE

ACCESSORY

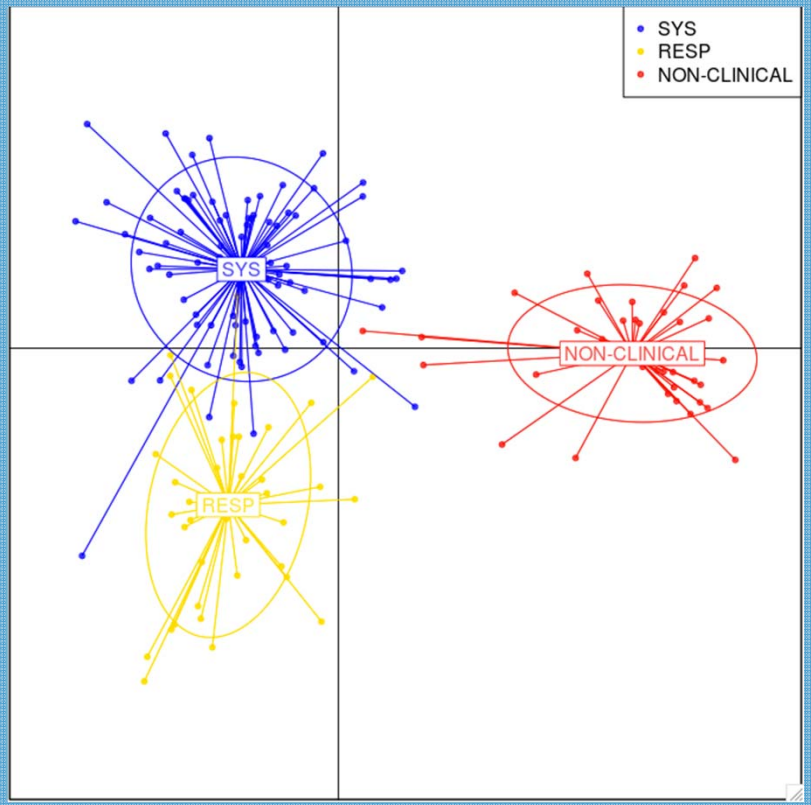


Results – Accessory Genome

PCA on the accessory genome



DAPC of the accessory genome



Results – Serotyping Markers

We have a shortlist of serovar–specific markers from our analyses of the capsule genes from HPS. From this and the accessory genome we have identified multiple statistically significant proteins for serovars 4, 5, 7, 13, 15.

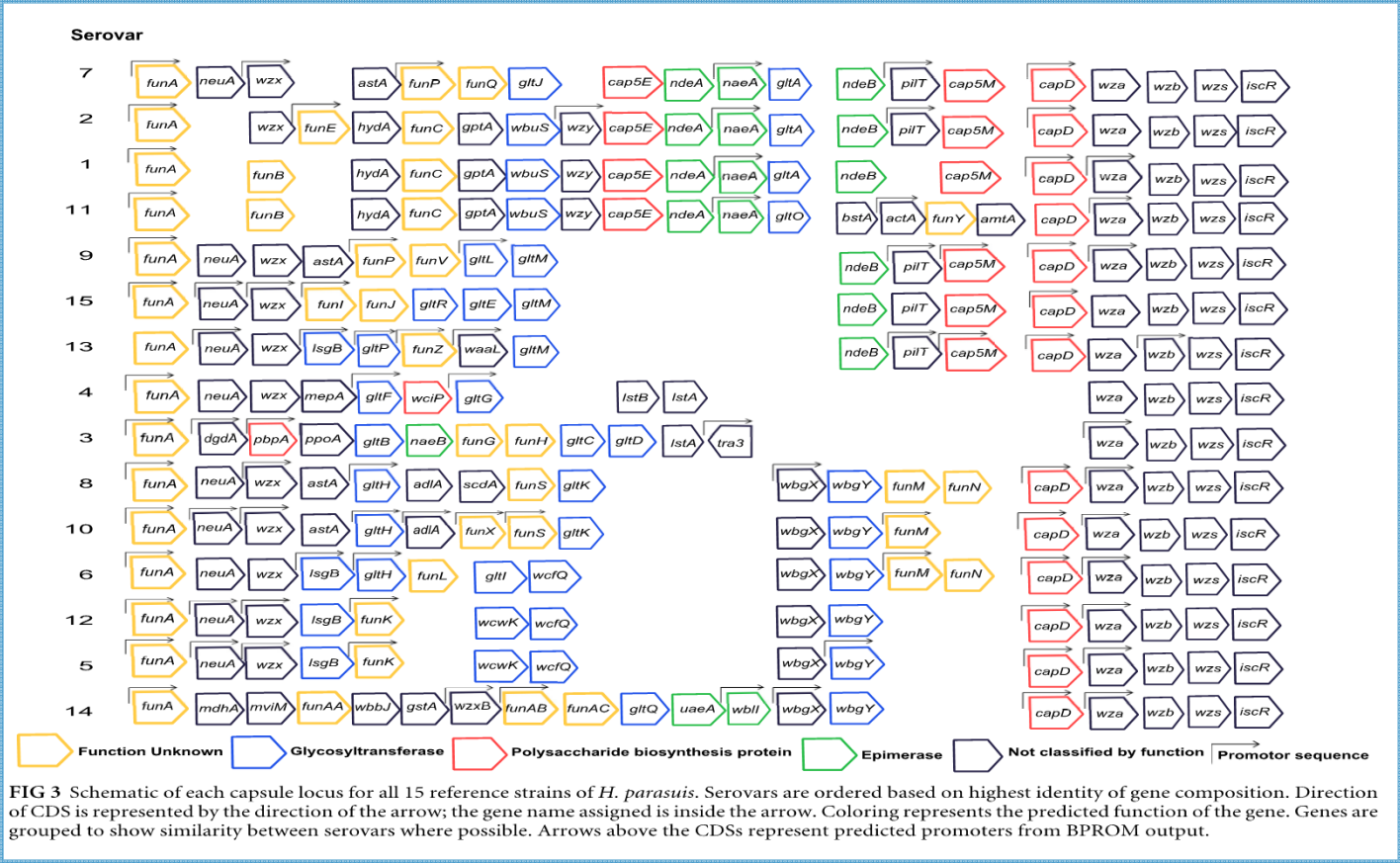


FIG 3 Schematic of each capsule locus for all 15 reference strains of *H. parasuis*. Serovars are ordered based on highest identity of gene composition. Direction of CDS is represented by the direction of the arrow; the gene name assigned is inside the arrow. Coloring represents the predicted function of the gene. Genes are grouped to show similarity between serovars where possible. Arrows above the CDSs represent predicted promoters from BPROM output.

Conclusions & Next Steps

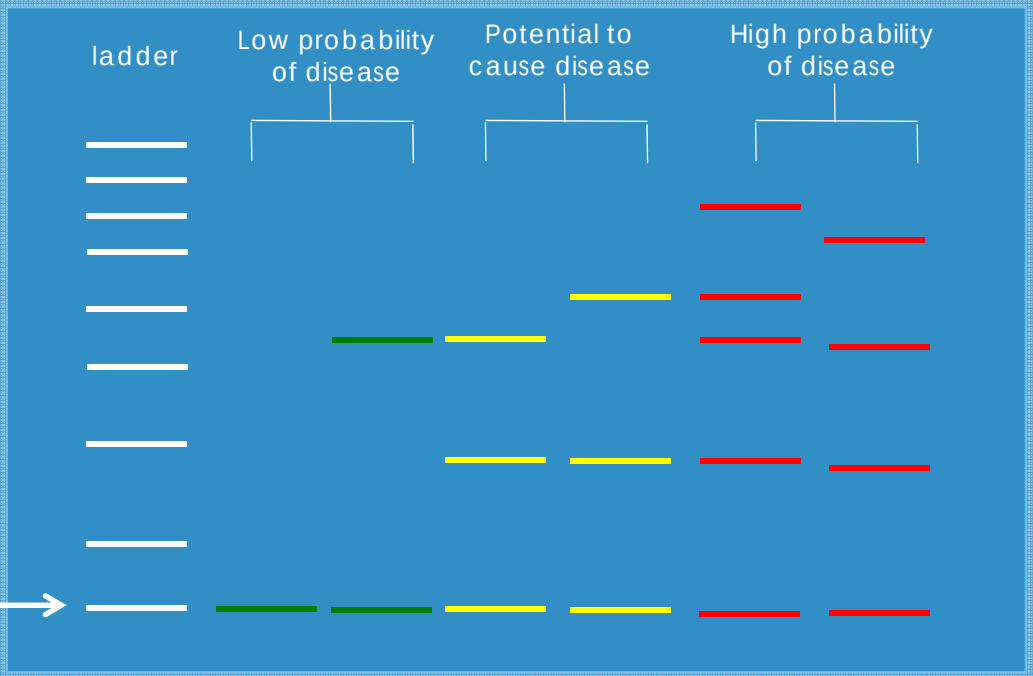
Work thus far has found statistically significant markers for serotype and for different disease categories for *H. parasuis*. The next stage is to determine which combination of markers gives the best probability of disease causation and to design the molecular diagnostics.

Preliminary Diagnostic Design

Representation of gel of ideal patho-typing assay able to predict the disease causing potential of a strain of each individual species with a traffic light system for predicting clinical importance of a strain

Markers could be able to predict:

- Disease
- Serotype
- Antibiotic Resistance



Acknowledgements

Thanks to:
BPEX
Dr. Lucy Weinert
Dr. Roy Chaudhuri
Dr. Shi-Lu. Luan
Dr. Sarah Peters
Dr Dan Tucker
Prof. Duncan Maskell

The Wellcome Trust Sanger Institute:
D. Harris
J. Parkhill

LOLA Project collaborators:
Prof. A. Rycroft
Prof. B. W. Wren
Prof. P. Langford

Strain contributors:
Dr. Virginia Aragon
Dr. Oystein Angen
AHVLA

References:

1.Fischer, S. *et al.* Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups. *Current protocols in bioinformatics* **Chapter 6,** Unit 6.12.1–19

2.Jombart, T. Ahmed, I. (2011) adegenet 1.3-1: new tools for the anlaysis of genome-wide SNP data. *Bioinformatics Applications Note*. Vol. 27 no.21 pg. 3070-3071.

3.Jombart, T., Devillard, S. & Balloux, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics* **11,** 94 (2010).

4.Olvera, A., Segalés, J. & Aragón, V. Update on the diagnosis of Haemophilus parasuis infection in pigs and novel genotyping methods. *The Veterinary Journal* **174,** 522–529 (2007).

5.Turni, C. & Blackall, P. J. Comparison of the indirect haemagglutination and gel diffusion test for serotyping Haemophilus parasuis. *Veterinary Microbiology* **106,** 145–51 (2005).

6.Zerbino, D. R. & Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* **18,** 821–9 (2008).

7.Olvera, A., Cerdà-Cuéllar, M. & Aragon, V. Study of the population structure of Haemophilus parasuis by multilocus sequence typing. *Microbiology* **152,** 3683–90 (2006).

8.Seemann, T. Prokka User Manual. (2013). at <http://bioinformatics.net.au/prokka-manual.html>

9.Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11,** 119 (2010).